

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problems Mailbox.**

**Best Available Copy**

## Pattern recognition

**Patent number:** CN1199488  
**Publication date:** 1998-11-18  
**Inventor:** DOWNEY S N (GB)  
**Applicant:** BRITISH TELECOMM (GB)  
**Classification:**  
- **International:** G10L3/00  
- **European:**  
**Application number:** CN19960197590 19960823  
**Priority number(s):** EP19950305982 19950824

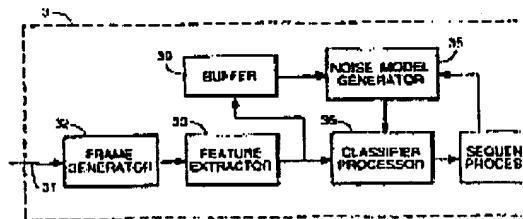
Also published as

WO97086  
US607881  
AU72051

Abstract not available for CN1199488

Abstract of correspondent: US6078884

PCT No. PCT/GB96/02069 Sec. 371 Date Mar. 26, 1998  
Sec. 102(e) Date Mar. 26, 1998 PCT Filed Aug. 23, 1998  
PCT Pub. No. WO97/08684 PCT Pub. Date Mar. 6, 1997  
Pattern recognition apparatus uses a recognition processor for processing an input signal to indicate its similarity to allowed sequences of reference patterns to be recognised. A speech recognition processor includes a classification arrangement to identify a sequence of patterns corresponding to said input signal and for repeatedly partitioning the input signal into a speech-containing portion and, preceding and/or following said speech-containing portion, noise or silence portions. A noise model generator is provided to generate a pattern of the noise or silence portion, for subsequent use by said classification means for pattern identification purposes. The noise model generator may generate a noise model for each noise portion of the input signal, which may be used to adapt the reference patterns.



CN5-01010-TA (1)

[19] 中华人民共和国专利局

[51] Int.Cl<sup>6</sup>  
G10L 3/00

## [12] 发明专利申请公开说明书

[21] 申请号 96197590.3

[43] 公开日 1998 年 11 月 18 日

[11] 公开号 CN 1199488A

[22] 申请日 96.8.23

[30] 优先权

[32] 95.8.24 [33] EP [31] 95305982.1

[86] 国际申请 PCT / GB96 / 02069 96.8.23

[87] 国际公布 WO97 / 08684 英 97.3.6

[89] 进入国家阶段日期 98.4.13

[71] 申请人 英国电讯公司

地址 英国英格兰伦敦

[72] 发明人 S · N · 道内

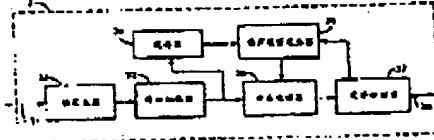
[14] 专利代理机构 中国专利代理(香港)有限公司  
代理人 程天正 张志醒

权利要求书 2 页 说明书 11 页 附图页数 7 页

[54] 发明名称 模式识别

[57] 摘要

模式识别装置包括一个用于处理一个输入信号以指出其与允许的基本模式序列的相似性的识别处理器 (3)。语音识别处理器 (3) 包括分类装置 (36、37)，所述分类装置确定对应于所说输入信号的模式序列，并重复地将所说输入信号划分成包含语音的部分和在所说语音部分之前和 / 或之后的噪声或静默部分。模式识别装置还有一个噪声模型发生器 (35)，用于产生对应于噪声或静默部分的模式，以便由所说的分类装置 (36、37) 在其后的模式识别操作中使用。该噪声模型发生器可以为所说输入信号的每个噪声部分产生一个噪声模型，该噪声模型可以用于自适应所说基本模式。



(BJ) 第 1456 号

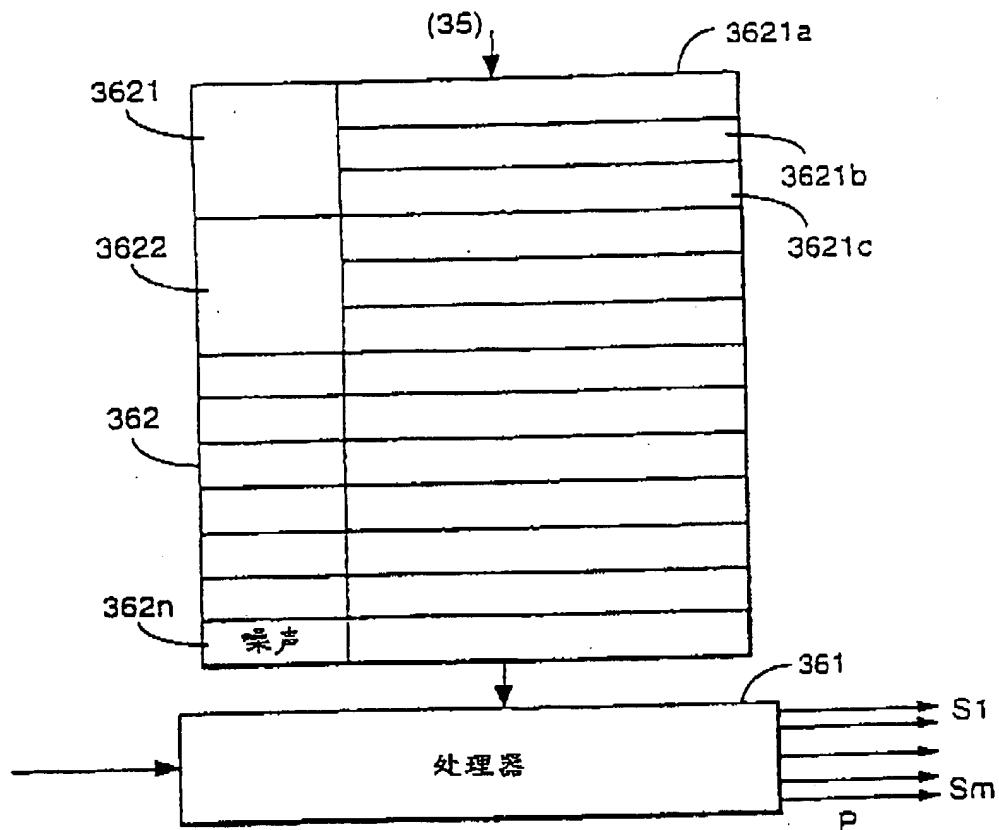


图 3

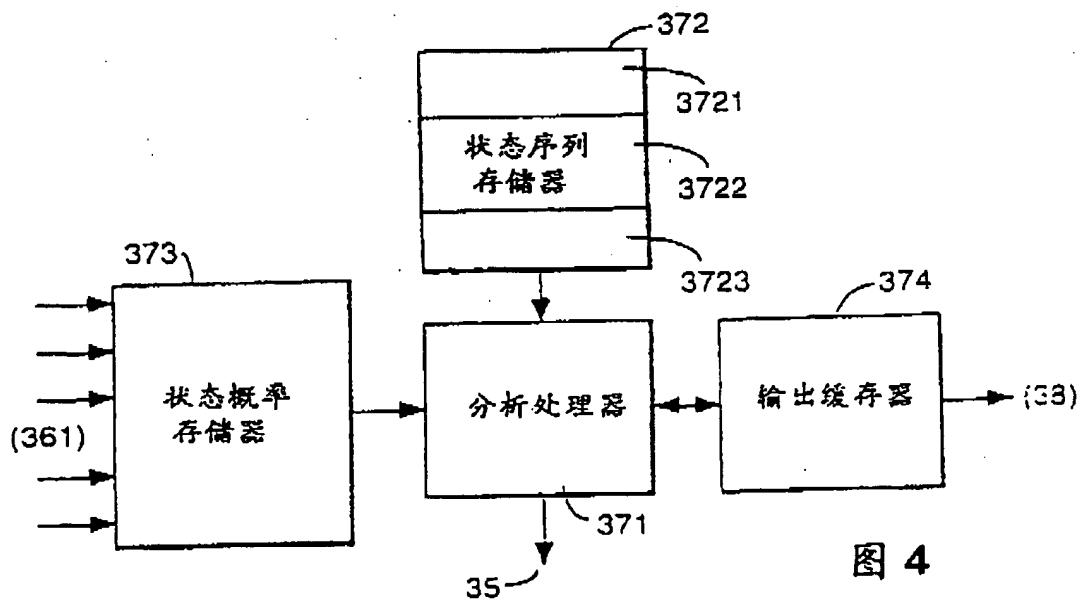


图 4

征系数。

最后，特征抽取器 33 输出对于每个连续帧加一的一个帧数。将所说特征向量输入分类器 36 和噪声模型发生器 35 中。FIFO 缓存器 39 在特征向量输入到噪声模型发生器 35 之前将它们缓存在其中。

5 在本实施例中帧发生器 32 和特征抽取器 33 是由一个适当编程的数字信号处理器 ( DSP ) ( 诸如 Motorola<sup>(TM)</sup> DSP 56000 , 或者 Texas<sup>(TM)</sup> Instruments TMS C320 ) 或者类似的装置构成的。

#### 分类器 36

参见图 3 , 在本实施例中, 分类器 36 包括一个分类处理器 361 和一个状态存储器 362 .

10 状态存储器 362 相对于每一个待识别的语音单位 ( 如音素变形 ) 包括一个状态信息组 3621 、 3622 . . . . . 例如, 由识别处理器识别的每一个音素变形由包括三种状态的一个隐藏马尔克夫模型表示, 因此为了存储每个音素变形的参数, 在状态存储器 362 中提供了三个状态信息组 3621a 、 15 3621b 、 3621c .

15 所说状态信息组存储限定表示相应音素变形的隐藏马尔克夫模型状态的参数, 这些参数已经按照常规方法根据一组学习数据来确定。状态存储器 362 还在一个状态信息组 362n 中存储构成平均线噪声估计值模型的参数, 所说估计值是按照常规方法随机生成的, 例如许多电话通话信号中生成。

20 对于每个输入其中的数据帧, 分类处理器 36 依次读取存储器 362 中的各个状态信息组, 并相对于各个数据帧, 利用当前输入的特征系数组来计算输入特征组或向量与相应状态对应的概率  $P_i$  。

因此, 分类处理器的输出是一组状态概率值  $P_i$  , 每个概率值对应于状态存储器 362 中的一个状态, 指出输入特征向量对应于各个状态的似然性。

25 所说分类处理器 361 可以是一个适当编程的数字信号处理器 ( DSP ) , 具体地说, 可以是与用作特征抽取器 33 相同的数字信号处理器。

#### 定序器 37

30 参见图 4 , 在本实施例中定序器 37 包括一个状态序列存储器 372 、一个分析处理器 371 、和一个定序器输出缓存器 374 .

定序器还包括一个状态概率存储器 373 , 其为每一个经过处理的数据帧

存储分类处理器 361 的输出。所说状态序列存储器 372 包括若干状态序列信息组 3721、3722、……，每个信息组对应于由音素变形和噪声信息串构成的、待识别的一个字或短语序列。

如图 5 所示，状态序列存储器 372 中的每个状态序列包括许多状态  $S_1$ 、  
5  $S_2$ 、……、 $S_N$  和相对于每个状态的两个概率值：重复概率( $P_{ii}$ )和朝向下一个状态的过渡概率( $P_{i,i+1}$ )。所说的序列状态就是多个三状态组，每一三种状态组与一个音素变形相关，在某些情况下，还与噪声相关。所以所观测的、与一系列数据帧相关联的状态序列可以由每个状态序列模型 372i 等元中的每个状态  $S_i$  的多次重复构成；例如：

10

帧数	1	2	3	4	5	6	7	8	9	…	Z	Z+1
状态	$S_1$	$S_1$	$S_1$	$S_2$	$S_2$	$S_2$	$S_2$	$S_2$	$S_2$	…	$S_n$	$S_n$

分析处理器 371 用于在每一数据帧期间读取存储在状态概率存储器 373 中的状态概率，计算状态与时间日期的最可能对应方式，并将该结果与存储在状态序列存储器 372 中的每一个状态序列相比较。例如，该状态序列可以包括电话簿中的姓名或数字串。

15

所说计算应用在上文引用的 Cox 的论文中所述的、众所周知的隐藏马尔克夫模型方法。可取的是，分析处理器 371 进行的隐藏马尔克夫模型处理是采用著名的 Viterbi 算法。例如，分析处理器 371 可以是一个微处理器如 Intel<sup>(TM)</sup> i-486<sup>(TM)</sup> 微处理器或 Motorola<sup>(TM)</sup> 68000 微处理器，或者还可以是一个 DSP 装置（例如，与用作前述处理器相同的 DSP 装置）。

20

所以，对于每个状态序列（对应于待识别的字、短语或其它语音序列），分析处理器 371 在输入语音的每一帧期间输出一个概率评分，并存储在输出缓存器 374 中。因此，存储器 374 包括关于每一帧输入信号和每一序列的一个概率评分、一个帧数记录和所说概率评分涉及的状态模型的记录。当检测到发音结尾时，从该缓存器向输出端口 38 输出一个指示最可能状态序列的标记信号，以指示已经识别的相应姓名、字或短语。

25

然后定序器处理器检查缓存器 374 中包含的信息，并借助于所说帧数，确定被识别为处于该语音识别装置的词汇表中的输入信号部分（这里称为语音部分）和被认为不处于所说词汇表中的输入信号部分（下文中称之为